

André M. Santamaría Regal

DNI: 36117001D

Enunciado

1.- Crea un archivo mapper.py que procese el archivo de la captura bluetooth que tienes en el aula virtual.

Debe funcionar de una manera muy parecida al contador de palabras con la modificación que no todas las líneas son significativas. La única que nos interesa en este problema es la línea en la que aparece la MAC capturada en la que procesaremos única y exclusivamente la MAC, excluyendo la etiqueta "Dirección:".

Como respuesta a esta pregunta muestra tu código comentado.

```
Mapper2.py U x
Curso de IA & Big Data > Big data aplicado > Practica Contador de Macs con MapReduce > Mapper2.py > ...
1  #!/usr/bin/python3
2
3  entrada = open("bluetoothLog.txt", "r", encoding= "utf8")
4  salida = open("salida_mapper.txt","w")
5  label = "Dirección:"
6  #Leemos cada linea del archivo
7  for line in entrada:
8      #si la linea contiene la label indicada divide el texto por dicha label
9      if label in line:
10         divide = line.split(label)
11         #comprobamos si la lista divide tiene mas de un elemento significa que hubo division
12         if len(divide) > 1:
13             #cojemos el segundo elemento que es la direccion mac que es lo que buscamos y se escribe en el archivo de salida
14             macs = divide[1].strip()
15             print (macs, "\t", 1)
16             salida.write(macs + "\t" + "1\n")
17
18  salida.close()
19
```


2.- (Opcional) Si tienes Hadoop funcionando no será necesario que pases por esta fase ya que se encargará Hadoop de hacerlo, puedes pasar al siguiente enunciado. Si tienes que ejecutar en local adapta tu código y aplica el archivo Python “ordenar.py” al archivo “salida_mapper.txt” para ordenar los resultados en un nuevo archivo “entrada_reducer.txt”. Muestra la salida del comando “head -n 20 entrada_reducer.txt” para ver las 20 primeras líneas del resultado.

3.- Crea un archivo llamado reducer.py nos devuelva el número de apariciones de cada MAC. Como respuesta muestra tu código comentado.

```
Mapper2.py U ordenar.py U reducer.py U X
Curso de IA & Big Data > Big data aplicado > Práctica Contador de Macs con MapReduce > reducer.py > ...

1  #!/usr/bin/env python
2  import sys
3
4  contador = 0
5  mac_anterior = None
6  entrada = open("entrada_reducer.txt", "r", encoding= "utf8")
7
8  for linea in entrada:
9      linea = linea.strip()
10     lista = linea.split("\t")
11     mac = lista[0]
12
13     if mac_anterior == None:
14         mac_anterior = mac
15
16     if mac == mac_anterior:
17         contador += 1
18
19     else:
20         print(mac_anterior, "\t", contador)
21         mac_anterior = mac
22         contador = 1
23
24     print(mac_anterior, "\t", contador)
25
```

PROBLEMS	OUTPUT	DEBUG CONSOLE	TERMINAL	PORTS
	FE:16:94:1D:75:E8		7	
	FE:2A:33:34:51:18		23	
	FE:40:07:08:E3:34		7	
	FE:40:F4:5D:41:6F		1	
	FE:51:AD:CD:8B:86		2	
	FE:52:35:41:3C:6A		5	
	FE:59:E5:6C:46:3C		13	
	FE:61:0D:15:19:81		4	
	FE:61:A1:8F:07:69		21	
	FE:6B:58:9B:7E:A4		37	
	FE:7B:38:12:58:27		7	
	FE:7D:8A:94:16:7D		6	
	FE:9A:43:49:C9:45		1	
	FE:9D:19:A3:C9:58		3	
	FE:B7:9A:8C:14:16		2	
	FE:B8:8D:7C:04:37		13	
	FE:C1:19:92:76:58		5	
	FE:CC:28:A8:9A:59		4	
	FE:DF:E1:C1:06:01		1	
	FE:E0:40:D2:17:DC		20	
	FE:E4:C6:94:CA:9A		1	
	FE:EB:A6:FD:8B:3C		6	
	FF:43:05:D0:8D:FC		70	
	FF:8E:8A:2C:1D:38		2	
	FF:14:99:7A:20:FC		3	
	FF:14:67:84:1A:65		22	
	FF:0D:95:01:E2:8E		3	
	FF:43:94:2F:24:71		10	
	FF:6A:B7:43:38:81		11	
	FF:76:47:0C:38:F6		38	
	FF:98:97:D5:AE:D4		8	
	FF:C5:AA:E8:3F:6C		70	
	FF:07:96:03:59:18		14	
	FF:0D:13:0A:79:E7		1	

PS C:\Users\dres\OneDrive\Documents\Curso de IA & Big Data\Big data aplicado\Practica Contador de Macs con MapReduce> []

4.- Prueba tu mapper y reducer en un clúster Hadoop con HDFS y YARN. Indica el comando que usas para ver los resultados y realiza una captura de pantalla de parte del resultado.